DEPARTMENT OF DEFENCE
DEFENCE SCIENCE & TECHNOLOGY ORGANISATION | DSTO

Feasibility and Genetic Algorithms: The Behaviour of Crossover and Mutation

Darryn J. Reid

DSTO-TN-0320

Feasibility and Genetic Algorithms: The Behaviour of Crossover and Mutation

Darryn J. Reid

Land Operations Division
Electronics and Surveillance Research Laboratory

DSTO-TN-0320

ABSTRACT

New genetic operators are described that assure preservation of the feasibility of candidate solutions to any discrete and linearly constrained optimisation problem. The design of these operators is the result of extensive theoretical investigations, with particular assiduity devoted to considering the most challenging examples of this type.

Attention is largely centred on problems that have defied satisfactory solution by traditional means, because of poorly behaved or imprecise objective functions, high dimensionality, or intractable algorithmic complexity. Evolutionary Algorithms, inspired by adaptation in biological systems, are ideally suited to such arduous conditions, as testified by their increasing popularity and successful exploitation in resolving numerous difficult problems.

Their obvious attraction notwithstanding, the applicability of evolutionary algorithms has suffered by the deficiency of general techniques to manage constraints, a feature common to many problems, and the asperity is compounded when some proportion of the decision variables are discrete. The enhanced operators presented here guarantee the feasibility of proffered aspirant solutions with respect to a system of linear constraints.

Exploration of highly constrained systems indicates the possibility of performance degradation, with the diminishing probability of discovering a feasible and meaningful exchange of information between candidate solutions. Relaxation of the crossover operator is proposed to alleviate this, and the effective utilisation of this in the context of the overall algorithm suggests a modification in which the population is permitted to transiently increase in size.

RELEASE LIMITATION

Approved for public release

DEPARTMENT OF DEFENCE
DEFENCE SCIENCE & TECHNOLOGY ORGANISATION DSTO

Published by

DSTO Electronics and Surveillance Research Laboratory
PO Box 1500
Salisbury South Australia 5108

Telephone: (08) 8259 5555
Fax: (08) 8259 6567

AR-011-641
December 2000

Conditions of Release and Disposal

Approved for Public Release

Feasibility and Genetic Algorithms: The Behaviour of Crossover and Mutation

EXECUTIVE SUMMARY

This document details the results of enabling research aimed at providing advanced techniques for building sophisticated automated planning and decision support tools. In particular, a broad range of apparently dissimilar resource management problems such as surveillance planning, resupply distribution planning, and fire support planning are, in fact, intimately related.

Solving these kinds of problems can be deceptively difficult; they may require unrealisable computational resources to resolve problems of even modest size. In addition, some of these planning and decision support functions display complicated interrelationships between the factors that together define the relative desirability of one possible solution over another. This nonlinearity itself poses substantial barriers to constructing effective tools for improving the timeliness and quality of decisions.

Additional considerations arise from the nature of the context of military operations and the translation of the military problem into mathematical terms; some subtle and poorly understood factors may not be fully represented in the mathematical formulation. The accepted resolution to this is to augment the ability of the human decision-maker with automated assistance, rather than to attempt to completely replace the human with a fully automated tool, at least in the first instance. To achieve this, the proposed solution technologies should provide the human with a range of plausible solutions to consider before he or she implements a final decision.

In response to these realisations, research has focussed on discrete optimisation problems and the discrete component of mixed optimisation problems that defy satisfactory elucidation by more traditional means. Attention is particularly focussed on problems of high dimensionality that describe poorly behaved objective functions, and those for which fully deterministic algorithms present intractable algorithmic complexity. It is also presumed that an appropriate method of solution should develop a range of qualitatively comparable candidate solutions close to the theoretical optima.

Evolutionary and genetic algorithms are a class of randomised search methods that are squarely aimed at problems of this character. Inspired by the adaptive behaviour of biological organisms, all evolutionary and genetic algorithms maintain a population of potential solutions, with each new generation developed from the last using a set of

stochastic transition operators. This contrasts sharply with traditional approaches, which typically produce a sequence of single solutions deterministically.

Despite their inherent robustness and obvious utility, the application of evolutionary and genetic algorithms to resource management problems suffers from their inability to adequately manage systems of constraints. This deficiency is compounded further with the need to constrain decision variables to integral values. This document firstly examines the behaviour of usual genetic crossover and mutation operators with respect to any system of constraints, by describing the probabilities that the generated solutions lie within the feasible region that such a constraint system defines. Building further upon this basis when the constraints are all linear, enhanced operators are proposed that guarantee the feasibility of their results.

Highly constrained problems are characterised as those that possess a very small feasible region with respect to its smallest encompassing interval. Investigation of the behaviour of genetic operators in such a setting reveals the potential for greatly degraded performance as the probability of generating only feasible candidate solutions collapses. However, the probability that at least some of the solution proffered by crossover and mutation may remain significant. This suggests a scheme in which the population is permitted to fluctuate in size.

Author

Darryn J. Reid
Land Operations Division

Darryn J Reid completed the BSc and BscHons degrees in mathematics and computer science from the University of Queensland in 1991 and 1992, respectively, and received a PhD from the University of Queensland in 1995. Darryn Reid joined DSTO as a Senior Professional Officer grade C in the Information Technology Division in 1995, and subsequently became a Research Scientist in 1996. He joined Electronic Warfare Division as a Senior Research Scientist in 1999, and transferred to Land Operations Division in 2000.

Contents

1. Introduction

The Evolutionary and Genetic Algorithms are a class of stochastic relaxation techniques, based on the progressive adaptation of biological systems, that are attracting increasing interest for the resolution of a wide variety of challenging applications [1-12]. In contrast to sequential search methods, a genetic algorithm maintains a large population of candidate solutions, rather than a single iterate, and each population is developed by the application of stochastic transition operators to its predecessor.

Genetic algorithms also differ in the character of the development of new points for exploration; while most traditional methods act directly on the value of an iterate, the genetic operators utilise an encoding of the solution, in the form of a string in some appropriate alphabet. This string is the 'chromosome' that defines an individual population member, as distinct from its actual value, or 'phenotype'.

These operators can be grouped according to their biologically-inspired functions: selection operators define the virility with which each population member reproduces, crossover operators describe the mechanism by which individuals mate to produce offspring, and mutation allows an occasional alteration in genetic structure. Historically, the term 'Evolutionary Algorithm' has related to methods that utilise selection and mutation alone, while 'Genetic Algorithms' also employ crossover, and sometimes a variety of other additional mechanisms. Here, the distinction will not be vigorously supported, as the discussion presented permits either.

The attraction of the genetic and evolutionary algorithms is enhanced by their conceptual simplicity, and the capacious account of disparate problems that have yielded to this kind of approach. Although its exploitation to the enhancement of overall performance is possible, no gradient or other additional information about the objective function need be available, so that noisy, stiff, and non-differentiable problems in principle pose no additional challenge. Furthermore, the maintenance of a large population of candidate solutions reduces the tendency to converge to local, but not global, optima, rendering this approach a vigorous contender in the illation of multimodal problems.

Despite their obvious allure, the applicability of genetic algorithms has suffered in the lack of adequate and sufficiently general techniques for the management of systems of constraints [3,6-11]. In most existing approaches, the genetic operators are permitted to generate infeasible candidate solutions, and these are penalised by modifications of the objective function (penalty functions), or later corrected (repair algorithms). The adoption of highly specialised representations of the problem (decoder algorithms),

that ensure, or magnify the probability of, the manufacture of feasible solutions, have also been suggested.

Penalty function methods [3,9-10] discriminate against a possible solution according to its violation of the constraint system. Regardless of the magnitude of these penalties, the result may be calamitous, particularly if the number of infeasible solutions generated in each iteration is large. Both decoder and repair algorithms are frequently discovered to be computationally intensive, difficult to design, and highly problem-specific [6,11]. Furthermore, a decoder that still permits the possibility of developing an infeasible solution necessitates the utilisation also of penalties or repair algorithms, enhancing further the undesirability of this approach.

In response to these realisations, awareness has more recently been focused [6-8] on the alteration of the genetic operators themselves to preserve feasibility. In doing so, the potential of the constraint system to reduce the size of the search space is fully realised, facilitating the possibility of greatly improved behaviour. The additional cost imposed by these modifications is more than offset by the would-be burden of penalty functions, decoders, or repair algorithms.

It is a new methodology, conforming to this philosophy, for the resolution of problems constrained to integer values and by a system of linear inequalities that is described here. However, it should be understood that this might be easily combined with methods for continuous variables presented elsewhere [6], for the solution of mixed-integer type models. Within this context, then, the aspiration is the minimisation or maximisation of some function

$$f : \mathbf{Z}^n \to \mathbf{R},$$

subject to the constraint system

$$A \cdot x \leq b,\ l \leq x \leq u,\ x \in \mathbf{Z}^n,$$

where $A = [a_{i,j}] \in \mathbf{R}^{m \times n}$ and $b = [b_i] \in \mathbf{R}^m$, with $l = [l_j] \in \mathbf{Z}^n$ and $u = [u_j] \in \mathbf{Z}^n$.

Note also that no equality constraints are included; these may be converted to a system of this form with a reduction in the number of variables, and therefore also in the size of the search space [6-8].

2. Two-Point Crossover

The genetic selection operator is the emblematisation of the process of natural selection in biological systems, in which individuals survive to reproduce at varying rates [1,3-5,7-8,11]. The genetic algorithm emulates this discrimination by constructing a 'mating

pool', with each individual represented according to its 'fitness', which is its objective function value compared in some sense with those of the rest of the population. Pairs of individuals will then be chosen at random from this mating pool to undergo crossover.

Modelled on the reproduction of biological organisms by the exchange of genetic information, the crossover operator facilitates the parturition of new individuals from old. Crossover speculates on new solutions constructed by combining fragments of old solutions that have previously worked well, according to their fitness values. The emergent population is therefore an innovation comprised of solutions that are the juxtaposition of fragments that have been successful in the past.

The two-point crossover achieves this by swapping the components of two vectors in some random interval. For each pair of parent solutions chosen from the mating pool, two randomly generated indices mark the boundaries of the corresponding segments of information to be inter-changed.

In seeking to understand the ramifications of restricting this operator to the development of feasible solutions, consider the feasible region, which is a convex polyhedron P in n-dimensional space,

$$P = \left\{ x \in \mathbf{Z}^n : A \cdot x \le b \wedge l \le x \le u \right\}.$$

This set is covered by the smallest bounded set $\Lambda \subset \mathbf{Z}^n$ that forms a closed interval around P,

$$\Lambda \equiv \Lambda(P) = \left\{ x \in \mathbf{Z}^n : \bar{l} \le x \le \bar{u} \right\},$$

where

$$\bar{l} = [\bar{l}_i]_{i=1}^n \bullet \bar{l}_i = \min_x \left\{ x_i : x = [x_j]_{j=1}^n \wedge A \cdot x \le b \wedge l \le x \le u \right\},$$

and

$$\bar{u} = [\bar{u}_i]_{i=1}^n \bullet \bar{u}_i = \max_x \left\{ x_i : x = [x_j]_{j=1}^n \wedge A \cdot x \le b \wedge l \le x \le u \right\},$$

are the vectors whose components are the smallest and largest, respectively, corresponding components of all feasible solutions x.

The crossover operator takes as its arguments two vectors in n-dimensional space and the indices defining the crossover segment, and produces from this a new pair of vectors. No order is defined between the vectors of such a pair, and therefore all such pairs is a set

$$\Omega = \left\{ \{x, y\} : x, y \in \mathbf{Z}^n \right\}.$$

The indices defining the information exchange indicate some number of consecutive corresponding components of the argument vectors; the set of all possible indices is described by the set

$$\chi = \{i : 1 \le i \le n \wedge i \in \mathbf{N}\}.$$

Two-point crossover is therefore a function

$$\Xi : \Omega \times \chi \times \chi \to \Omega$$

defined as

$$\Xi(\{x, y\}, h, k) = \begin{cases} \{x', y'\} & \text{if } h < k \\ \{x, y\} & \text{otherwise} \end{cases}$$

and where x, y, and the resultant vectors x', y' have the form

$$x = [x_1 \cdots x_n]^T,\ y = [y_1 \cdots y_n]^T,$$

$$x' = [x_1 \cdots x_h\ y_{h+1} \cdots y_k\ x_{k+1} \cdots x_n]^T,$$

$$y' = [y_1 \cdots y_h\ x_{h+1} \cdots x_k\ y_{k+1} \cdots y_n]^T.$$

The crossover is called 'meaningful' if $h < k$; when $h \ge k$ no exchange of information occurs. Furthermore, the set of application of crossover will, for convenience, be abbreviated as follows:

$$\Xi_{h,k}(x, y) \equiv \Xi(\{x, y\}, h, k).$$

Any meaningful crossover for which $\Xi_{h,k}(x, y) = \{x, y\}$ will be termed 'trivial', including the case when $x = y$. All other meaningful crossovers, having the potential of generating points previously unseen, are 'non-trivial'.

Clearly, the crossover of feasible solutions cannot result in the production of points outside the surrounding interval $\Lambda \subset \mathbf{Z}^n$ (note here also that the converse is not true). Therefore, in considering the result of a crossover operation, attention may be entirely restricted to this interval. In the presumption that the points forming the old population are all feasible, the crossover will only be applied to elements of the feasible region P. That is, only crossovers of the form

$$\Xi_{h,k}(x, y) \bullet (\{x, y\}, h, k) \in \Phi(P)$$

are relevant, where $\Phi(P)$ is the set of all meaningful crossovers between points in any set S such that $S \subseteq \Lambda$,

$$\Phi(S) = \{(\{x,y\},h,k) : x,y \in S \wedge x \neq y \wedge h,k \in \chi \wedge h < k\}.$$

The number of meaningful crossovers between the points of a set $S \subseteq \Lambda$ is easily derived [7].

Lemma 2.1 [7]:

The number of meaningful crossovers between solutions in $S \subseteq \Lambda$ is

$$|\Phi(S)| = \frac{1}{4}(n^2 - n) \cdot |S| \cdot (|S| - 1).$$

In particular, when $S = P$, this yields the number of meaningful crossovers possible between feasible points.

The potential exists for the extemporaneous generation of infeasible candidate solutions, whenever the number of points $|\Lambda|$ in the enclosing interval Λ transcends the number $|P|$ in the feasible region P. The design of an appropriate mechanism for managing the crossover depends upon exactly how great this propensity is.

The most immediate strategy is to permit only those crossovers that would manufacture nothing but feasible points; crossovers of this type, namely

$$\{(\{x,y\},h,k) \in \Phi(P) \wedge \Xi_{h,k}(x,y) \subseteq P\},$$

will be termed 'feasible', or 'entirely feasible' in emphasis of the fact that only feasible solutions are manufactured.

The number of crossovers that produce only feasible points, from points in P, potentially bears considerable impact on the overall performance of the genetic algorithm. If the likelihood of a successful search for a crossover segment that preserves feasibility is remote, then little progress towards convergence is possible.

Theorem 2.2 [7]:

The probability that a crossover randomly chosen from $\Phi(P)$ is entirely feasible is

$$\Pr\left(\Xi_{h,k}(x,y) \subseteq P\right) = \frac{|P| \cdot (|P| - 1)}{|\Lambda| \cdot (|\Lambda| - 1)}.$$

A graph of this probability distribution appears in [8]. Put another way, of the $\frac{1}{4}(n^2-n)\cdot|P|\cdot(|P|-1)$ meaningful crossovers between feasible points,

$$\begin{aligned}\overline{M}(P) &= \Pr\left(\Xi_{h,k}(x,y)\subseteq P\right)|\Phi(P)| \\ &= \frac{1}{4}(n^2-n)\cdot\frac{|P|^2\cdot(|P|-1)^2}{|\Lambda|\cdot(|\Lambda|-1)}\end{aligned} \tag{1}$$

of them are entirely feasible.

As the size of the feasible region P diminishes with respect to that of the enclosing interval Λ, so too does the probability (Theorem 2.2) of a crossover being feasible. The result is a decrease in the number of entirely feasible crossover operations (see equation (1)), and consequently in the performance of the genetic algorithm overall.

The amelioration of this is facilitated in the realisation that not all crossovers producing feasible solutions are themselves entirely feasible. The class of crossovers that produce only feasible solutions is contained within the wider set of crossovers that produce at least one feasible point. By exploiting this, at least one new solution might be obtained, where none could be illuminated before.

The set of 'half-feasible' crossovers, that produce only one feasible point, is

$$\left\{(\{x,y\},h,k)\in\Phi(P):|\Xi_{h,k}(x,y)\cap P|=1\right\},$$

and its size can be ascertained similarly to that of the entirely feasible crossovers, by first considering the probability that a crossover between points in P belongs to this set.

Theorem 2.3 [7]:

The probability that a crossover randomly chosen from $\Phi(P)$ is half-feasible is

$$\Pr\left(|\Xi_{h,k}(x,y)\cap P|=1\right)=\frac{2\cdot|P|\cdot(|\Lambda|-|P|)}{|\Lambda|\cdot(|\Lambda|-1)}.$$

A graph of this probability distribution can be found in [8]. The number of half-feasible crossover operations, with respect to the feasible region P, is therefore

$$\overline{H}(P)=\frac{1}{2}(n^2-n)\frac{|P|^2\cdot(|P|-1)\cdot(|\Lambda|-|P|)}{|\Lambda|\cdot(|\Lambda|-1)}. \tag{2}$$

Comparing Equations (1) and (2), the number of half-feasible crossovers is discovered to surpass that of the entirely feasible variety, when the feasible region consumes less than about two-thirds of its enveloping interval:

$$\overline{H}(P) > \overline{M}(P) \text{ whenever } |P| < \frac{2}{3}(|\Lambda|+1)\,. \tag{3}$$

Under such conditions, the utilisation of the half-feasible crossovers may provide considerable advantages over restricting sedulity to those that produce only feasible solutions. However, the realisation of this potential to the greatest advantage itself demands careful application.

3. Maintaining Feasibility

In order to afford the limitation of the choice of crossover segment to those that produce only feasible solutions, a sensible and efficient search is required. The exploration of all potential crossovers, for a given pair of points, is prohibitively expensive; first appointing one of the segment boundaries, by a random selection, reduces the number of possibilities from $O(n^2)$ to $O(n)$.

The actual search itself might then be effectuated by producing, in turn, each of the would-be resultant individuals, and appraising their satisfaction of the constraint system. However, in the case of a linear system, as prescribed, this disregards the relationship that exists between crossovers that have segments with one boundary in common. More thorough excogitation reveals that the entire search for an entirely feasible crossover between two given vectors can be made in a total of $O(mn + n^2)$ time.

For brevity of notation, the residual $r : \mathbf{Z}^n \to \mathbf{R}^m$ of a vector in n-dimensional integer space is a function defined as

$$r(x) = b - A \cdot x\,,$$

representing the amount by which a solution vector x violates or satisfies the constraints. The proposal is made that the residual of each vector be stored with it, to enable a rapid search for feasible crossovers; the residuals of the new vectors produced can be quickly found from those of the parents. The search method itself is justified by the following theorem.

Theorem 3.1 [7]:

Consider the crossover $\{x', y'\} = \Xi_{h,k}(x, y)$, with x' and y' as described in the definition of crossover presented in Section 2. Then

$$x' \in P \Leftrightarrow \delta^{(h,k)} \le r(x)\text{, and } y' \in P \Leftrightarrow -r(x) \le \delta^{(h,k)},$$

where $\delta^{(h,k)} \equiv \delta^{(h,k)}(x, y)$ satisfies

$$\delta^{(h-1,k)} = \delta^{(h,k)} + (y_h - x_h) \cdot [a_{i,h}]_{i=1}^{m},$$

$$\delta^{(h,k+1)} = \delta^{(h,k)} + (y_{k+1} - x_{k+1}) \cdot [a_{i,k+1}]_{i=1}^{m},$$

$$\delta^{(h,h)} = 0,$$

for any $h,k : 1 \le h \le k \le n$. Furthermore,

$$r(x') = r(x) - \delta^{(h,k)} \text{ and } r(y') = r(y) + \delta^{(h,k)}.$$

Given a pair of vectors from the mating pool, the quest for a feasible crossover begins with the random choice of the first segment boundary. This may be interpreted either as the lower boundary h, or as the upper k, so the search proceeds in two stages. All feasible crossovers that have this chosen point as the lower delimitation h might be generated by initialising

$$\delta^{(h,k)} = 0$$

and successively applying the recurrence relation

$$\delta^{(h,k+1)} = \delta^{(h,k)} + (y_{k+1} - x_{k+1}) \cdot [a_{i,k+1}]_{i=1}^{m},$$

with each such crossover segment producing only feasible candidate solutions if

$$-r(y) \le \delta^{(h,k)} \le r(x).$$

A small example illustrating this procedure with a system of three linear constraints in five dimensions appears in [8].

The search, applying instead the alternative recurrence stated in the theorem, for crossovers with the chosen point used as the upper boundary instead, may be then performed. Early termination of the whole procedure can be accomplished by probabilistically deciding to accept or reject each feasible crossover, as they are illuminated. Given that there are a total of $n-1$ possible choices of crossover segments

that have any particular index as one of the boundaries, the probability that any random crossover in $\Phi(P)$ is entirely feasible (Theorem 2.2), and presuming a uniform distribution in the choice of the crossover from all those discovered to be feasible, the probability (on average) that a feasible crossover is chosen and applied [7] is

$$p_F = \frac{1}{(n-1) \cdot \Pr\left(\Xi_{h,k}(x,y) \subseteq P\right)}.$$

This assumes also that the expected number of crossovers found is at least one; when this condition is violated, the feasible region is exceedingly small, demanding instead the exploitation of a modified scheme, to be discussed later.

In most practical situations, the parameter p_F cannot be known prior to the application of the genetic algorithm, as it is defined in terms of the size of the feasible set. By keeping a count of the total number of feasible crossovers found t_F, and the total number of possibilities explored t_C, however, p_F can be approximated using

$$\Pr\left(\Xi_{h,k}(x,y) \subseteq P\right) \approx \frac{t_F}{t_C}.$$

The estimation of p_F used in the algorithm is periodically updated, typically at the end of each generation (major iteration). As the algorithm progresses, the greater the number of crossovers it samples, so that this approximation rapidly converges to the actual value.

Given the dominance of crossovers that manufacture a single feasible point over those that produce two, in more highly constrained examples (refer to equation (3)), the exploitation of these half-feasible crossovers must be considered when dealing with these kinds of problems. This undertaking requires the modification of the search procedure so far discussed, to recognise those crossovers for which only one of the resultant vectors is feasible.

This might be actualised by resorting to a half-feasible crossover only in the absence of any that are entirely feasible. Such a scheme respects a preference for the entirely feasible variety, but this may be at the expense of an early termination; when no feasible crossover is found, the search must complete before allowing the acceptance of a half-feasible crossover. This method appears most useful, then, when the number of entirely feasible crossovers in $\Phi(P)$ is reasonably high, yet the chance of discovering none cannot be ignored either, say when $|P| \approx \frac{2}{3}(|\Lambda|+1)$. However, this region may be quite narrow [7].

An alternative approach, most appropriate for the most highly-constrained examples (when $|P| << \frac{2}{3}(|\Lambda|+1)$), is to regard feasible and half-feasible crossovers uniformly, accepting either when unearthed with probability

$$p_{F+H} = \frac{1}{(n-1)\cdot\left[\Pr\left(\Xi_{h,k}(x,y)\subseteq P\right)+\Pr\left(|\Xi_{h,k}(x,y)\cap P|=1\right)\right]}.$$

This expression is derived using Theorems 2.2 and 2.3, similarly to that for p_F above, and can be likewise approximated by keeping a count t_{F+H} of the total number of entirely feasible and half-feasible crossovers discovered to date.

With the restriction that only feasible points are permitted to enter the new population, one of the solutions offered by the half-feasible crossover must be discarded. Replacing the infeasible offspring with either of the two parents is unacceptable, as this results in a forfeit of genotypic information; the values common to the discarded parent and infeasible child are lost to the algorithm.

This amounts to a pre-emptive form of selection, at the micro-scale, for a judgment is passed on some of the solution fragments involved, and without regard to the specifics of the overall selection mechanism being applied. By sustaining both the parent individuals and the new solution produced by the half-feasible crossover, the genotypic diversity of the population is maintained. Solutions are thereby appraised solely by the selection operator, in complete assent with its intended purpose. This requires a flexible population, permitted to grow evanescently in size up to half as much again, until selection reduces this in creating the next mating pool.

4. Feasible Mutation

The combined influence of selection and crossover, together furnishing most of the genetic algorithm's information processing efficaciousness, surmounts that of mutation (in contrast, the pure evolutionary strategies [1-2,4-5] rely solely or preponderately on the mutation operator). However, with the utilisation of selection and crossover alone comes the possibility of losing the ability to investigate some potentially beneficial slices of the search space.

A detrimental loss of knowledge might occur, for example, when some substructure of the optimum occurs in combination with other, overwhelmingly poor, partial solutions. Over-exuberance of the genetic algorithm in its quest may immediately eliminate this individual, and the desired substructure with it, from future consideration. Indeed, it may occur that no individual exists in the initial population that has some particular component value; perhaps this is a value that is required to construct an optimal (or a satisfactory near-optimal) solution.

The mutation operator, like its biological inspiration, implies a random alteration to some vector component (allele), occurring with low probability. Through mutation, an individual may become not viable; in biological systems this results in the individual's demise. In the confines of the artificial population, to permit the generation of infeasible solutions and then eliminate them is a waste of effort, particularly in problems where the feasible region and therefore the chances of retaining viability in an unrestrained and capricious dissimilation are small.

Mutation, like the crossover, should be restricted in respect for the constraint system. Towards this goal, consider that mutation is a function

$$\Psi : \mathbf{Z}^n \times \mathbf{Z} \times \chi \to \mathbf{Z}^n$$

defined as

$$\Psi\left([x_1 \cdots x_n]^T, v, k\right) = [x_1 \cdots x_{k-1}\ v\ x_{k+1} \cdots x_n]^T .$$

The mutation operator may also be viewed as a degeneration of the crossover operator; effectively, the mutated individual is crossed with a random vector, where the crossover segment is simply the single chosen allele. For this reason, the search for a feasible mutation takes a similar form to that for crossover.

Theorem 4.1 [7]:

The mutation

$$x' = \Psi(x, \varepsilon + x_k, k)\text{, where } \bar{l}_k - x_k \le \varepsilon \le \bar{u}_k - x_k \text{ and } \varepsilon \in \mathbf{Z},$$

is feasible if and only if, for all $i : 1 \le i \le m$,

$$a_{i,k} = 0 \vee \left(a_{i,k} > 0 \wedge \varepsilon \le \frac{r(x)_i}{a_{i,k}} \right) \vee \left(a_{i,k} < 0 \wedge \varepsilon \ge \frac{r(x)_i}{a_{i,k}} \right).$$

Applying this result is simple; given the solution vector and index of the component to be mutated, the interval of values of ε that would preserve feasibility can be determined, by considering in turn each element of the k th column of the constraint matrix A. A negative entry $a_{i,k}$ defines $r(x)_i / a_{i,k}$ as a lower bound on ε, and a positive entry describes this instead as an upper bound.

The possible mutation values so determined, the mutation itself can now be performed, by substituting the new value $x_k + \varepsilon$ for the old x_k. The calling function is

also free to decide whether to permit a trivial mutation (when $\varepsilon = 0$), or to try another component, or a different candidate solution.

5. Conclusion

The resolution by an evolutionary algorithm of a mathematical optimisation problem that is constrained to discrete values, and by a system of linear inequalities, has been disencumbered by the proposal of a set of new genetic operators that ensure the continued feasibility of candidate solutions. This is the emanation of theoretical investigations into the behaviour of the traditional two-point crossover operator, in the context of a set of viable solutions enclosed by some bounding interval.

Possible crossover operations applied to points of this feasible set are found to fall into two relevant categories: those that exclusively manufacture feasible solutions, and those that produce one point that is feasible, and one that is not. The proportions of these are understood to vary with respect to the character of the feasible region itself.

Upon this basis, a crossover operator is propounded that restricts itself to producing nothing but feasible candidates. However, this may undesirably limit the information-processing efficacy of the genetic algorithm, particularly when dealing with problems that are highly constrained, for which the number of crossovers solely yielding feasible solutions is greatly diminished. Techniques that utilise the half-feasible variety have been developed to overcome this, and their effective application alludes furthermore to a ephemeral increase in population size. Finally, a mutation operator is offered that likewise guarantees the feasibility of its result.

6. References

[1] Thomas Bach and Frank Hoffmeister. Extended selection mechanisms in genetic algorithms. *Proceedings of the Fourth International Conference on Genetic Algorithms*, pages 92-99, 1991.

[2] D B Fogel. Applying evolutionary programming to selected control problems. *Computers and Mathematics with Applications*, 27(11):89-104, 1994.

[3] David E Goldberg. *Genetic Algorithms in Search, Optimization, and Machine Learning*. Addison-Wesley, Reading, Massachusetts, 1989.

[4] David E Goldberg. Genetic and evolutionary algorithms come of age. *Communications of the ACM*, 37(3):113-119, March 1994.

[5] John Holland. Genetic algorithms. *Scientific American*, July 1992.

[6] Zbigniew Michalewicz and Cezary Z Janikow. Handling constraints in genetic algorithms. *Proceedings of the Fourth International Conference on Genetic Algorithms,* pages 151-157, 1991.

[7] Darryn J Reid. Using genetic algorithms in constrained discrete optimisation. *Mathematical and Computer Modelling,* 23(5) pp87-111, 1996.

[8] Darryn J Reid. Enhanced genetic operators for the resolution of discrete constrained optimization problems. *Computers and Operations Research,* 24(5) pp399-411, 1997.

[9] J T Richardson, M R Palmer, G Liepins, and M Hilliard. Some guidelines for genetic algorithms with penalty functions. *Proceedings of the Third International Conference on Genetic Algorithms,* pages 191-197, 1989.

[10] W Siedlecki and J Sklanski. Constrained genetic optimisation via dynamic reward-penalty balancing and its use in pattern recognition. *Proceedings of the Third International Conference on Genetic Algorithms,* pages 141-150, 1989.

[11] G A Vignaux and Z Michalewicz. A genetic algorithm for the linear transportation problem. *IEEE Transactions on Systems, Man, and Cybernetics,* 21(2), 1991.

[12] Darrell Whitley. The genitor algorithm and selection pressure: Why rank-based allocation of reproductive trials is best. *Proceedings of the Third International Conference on Genetic Algorithms,* pages 116-121, 1989.

DISTRIBUTION LIST

Feasibility and Genetic Algorithms: The Behaviour of Crossover and Mutation

Darryn J Reid

AUSTRALIA

DEFENCE ORGANISATION

Task Sponsor
DGC3ID

S&T Program
Chief Defence Scientist
FAS Science Policy } shared copy
AS Science Corporate Management
Director General Science Policy Development
Counsellor Defence Science, London (Doc Data Sheet)
Counsellor Defence Science, Washington (Doc Data Sheet)
Scientific Adviser to MRDC Thailand (Doc Data Sheet
Scientific Adviser Policy and Command
Navy Scientific Adviser (Doc Data Sheet and distribution list only)

Scientific Adviser - Army (Doc Data Sheet and distribution list only)

Air Force Scientific Adviser
Director Trials

Aeronautical and Maritime Research Laboratory
Director

Electronics and Surveillance Research Laboratory
Director (Doc Data Sheet and distribution list only)

Chief of Land Operations Division
Research Leader Land Systems
Bob Seymour
Darryn Reid

DSTO Library and Archives
Library Fishermans Bend (Doc Data Sheet)
Library Maribyrnong (Doc Data Sheet)
Library Salisbury (1 copy)
Australian Archives
Library, MOD, Pyrmont (Doc Data sheet only)

US Defense Technical Information Center, 2 copies

UK Defence Research Information Centre, 2 copies
Canada Defence Scientific Information Service, 1 copy
NZ Defence Information Centre, 1 copy
National Library of Australia, 1 copy

Capability Systems Staff
Director General Maritime Development (Doc Data Sheet only)
Director General Land Development
Director General Aerospace Development (Doc Data Sheet only)

Army
ASNSO ABCA, Puckapunyal, (4 copies)
SO (Science), DJFHQ(L), MILPO Enoggera, Queensland 4051 (Doc Data Sheet only)
NAPOC QWG Engineer NBCD c/- DENGRS-A, HQ Engineer Centre Liverpool Military Area, NSW 2174 (Doc Data Sheet relating to NBCD matters only)

Intelligence Program
DGSTA Defence Intelligence Organisation
Manager, Information Centre, Defence Intelligence Organisation

Corporate Support Program
Library Manager, DLS-Canberra

UNIVERSITIES AND COLLEGES
Australian Defence Force Academy
Library
Head of Aerospace and Mechanical Engineering
Hargrave Library, Monash University (Doc Data Sheet only)
Librarian, Flinders University

OTHER ORGANISATIONS
NASA (Canberra)
AusInfo
State Library of South Australia
Parliamentary Library, South Australia

OUTSIDE AUSTRALIA

ABSTRACTING AND INFORMATION ORGANISATIONS
Library, Chemical Abstracts Reference Service
Engineering Societies Library, US
Materials Information, Cambridge Scientific Abstracts, US
Documents Librarian, The Center for Research Libraries, US

INFORMATION EXCHANGE AGREEMENT PARTNERS

Acquisitions Unit, Science Reference and Information Service, UK

Library - Exchange Desk, National Institute of Standards and Technology, US

SPARES (5 copies)

Total number of copies: 45

DEFENCE SCIENCE AND TECHNOLOGY ORGANISATION
DOCUMENT CONTROL DATA

1. PRIVACY MARKING/CAVEAT (OF DOCUMENT)

2. TITLE	3. SECURITY CLASSIFICATION (FOR UNCLASSIFIED REPORTS THAT ARE LIMITED RELEASE USE (L) NEXT TO DOCUMENT CLASSIFICATION)
Feasibility and Genetic Algorithms: The Behaviour of Crossover and Mutation	Document (U) Title (U) Abstract (U)

4. AUTHOR(S)	5. CORPORATE AUTHOR
Darryn J. Reid	Electronics and Surveillance Research Laboratory PO Box 1500 Salisbury SA 5108

6a. DSTO NUMBER	6b. AR NUMBER	6c. TYPE OF REPORT	7. DOCUMENT DATE
DSTO-TN-0320	AR-011-641	Technical Note	December 2000

8. FILE NUMBER	9. TASK NUMBER	10. TASK SPONSOR	11. NO. OF PAGES	12. NO. OF REFERENCES
D 9505-19-135	ARM 98/118	DGC3ID	15	12

13. URL ON WORLDWIDE WEB	14. RELEASE AUTHORITY
http://www.dsto.defence.gov.au/corporate/reports/DSTO-TN-0320.pdf	Chief, Land Operations Division

15. SECONDARY RELEASE STATEMENT OF THIS DOCUMENT

Approved for public release

OVERSEAS ENQUIRIES OUTSIDE STATED LIMITATIONS SHOULD BE REFERRED TO DOCUMENT EXCHANGE, PO BOX 1500, SALISBURY, SA 5108, AUSTRALIA

16. DELIBERATE ANNOUNCEMENT

17. CASUAL ANNOUNCEMENT Yes

18. DEFTEST DESCRIPTORS

Genetic algorithm, Evolutionary algorithm, Linearly constrained optimisation.

19. ABSTRACT
New genetic operators are described that assure preservation of the feasibility of candidate solutions to any discrete and linearly constrained optimisation problem. The design of these operators is the result of extensive theoretical investigations, with particular assiduity devoted to considering the most challenging examples of this type.
Attention is largely centred on problems that have defied satisfactory solution by traditional means, because of poorly behaved or imprecise objective functions, high dimensionality, or intractable algorithmic complexity. Evolutionary Algorithms, inspired by adaptation in biological systems, are ideally suited to such arduous conditions, as testified by their increasing popularity and successful exploitation in resolving numerous difficult problems.
Their obvious attraction notwithstanding, the applicability of evolutionary algorithms has suffered by the deficiency of general techniques to manage constraints, a feature common to many problems, and the asperity is compounded when some proportion of the decision variables are discrete. The enhanced operators presented here guarantee the feasibility of proffered aspirant solutions with respect to a system of linear constraints.
Exploration of highly constrained systems indicates the possibility of performance degradation, with the diminishing probability of discovering a feasible and meaningful exchange of information between candidate solutions. Relaxation of the crossover operator is proposed to alleviate this, and the effective utilisation of this in the context of the overall algorithm suggests a modification in which the population is permitted to transiently increase in size.